

RESULT 11

US-10-072-349-234

: Sequence 234, Application US/10072349
: Publication No. US20030054420A1

: GENERAL INFORMATION:

: APPLICANT: Rosen et al.

: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

: FILE REFERENCE: P111001

: CURRENT APPLICATION NUMBER: US/10/072,349

: PRIOR APPLICATION REMOVED - See file wrapper or Palm

: NUMBER OF SEQ ID NOS: 334

: SOFTWARE: PatentIn Ver. 3.1

: SEQ ID NO 234

: LENGTH: 7221

: TYPE: DNA

: ORGANISM: Homo sapiens

US-10-072-349-234

Query Match 9.4%; Score 30.2; DB 14; Length 7221;

Best Local Similarity 48.0%; Pred. No. 19;

Matches 86; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 76 CCGTGAACCAATGGCGAGTCGCCCTACTATTCACACCACTTATTCAGAAAAGCGA 135

DB 6484 CTTCAAAAAGAACCTGATGTCACAAATCTACAAATCTAGATTCTAGATAGTGTG 6543

QY 136 CCGCGTCGATGATGATGCAATCTACCTCCACCAACCCAGCTTTGATCTGCTG 135

DB 6544 CTTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6603

QY 196 TGATCAGCAAGATGTCATACATGTCATATGATGATGATGATGATGATGATGATG 254

DB 6604 TCAAAAATTAAGATTAATTTTTCATCTACTACAGATATGATGATGATGATGATG 6662

US-10-050-704-57/c

: Sequence 57, Application US/10050704

: Publication No. US20030050442A1

: GENERAL INFORMATION:

: APPLICANT: Ruben et al.

: TITLE OF INVENTION: 62 Human Secreted Proteins

: FILE REFERENCE: P2039P1

: CURRENT APPLICATION NUMBER: US/10/050,704

: PRIOR APPLICATION REMOVED - See file wrapper or Palm

: NUMBER OF SEQ ID NOS: 344

: SOFTWARE: PatentIn Ver. 2.0

: SEQ ID NO 57

: TYPE: DNA

: ORGANISM: Homo sapiens

US-10-050-704-57

Query Match 9.3%; Score 29.8; DB 14; Length 1769;

Best Local Similarity 51.1%; Pred. No. 13;

Matches 70; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 152 TGCAATCTACTCTCAACCAACCCAGCTTTGATCTGCTTACTGATCAACCAAGCTTC 211

DB 1395 TGAACCT 1336

QY 212 CCTGATACGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 271

DB 1335 GCAGCTACGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276

QY 272 AAGAGAGAGGTTGGTCA 288

DB 1275 TGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1259

RESULT 13

US-09-766-399-42

: Sequence 43, Application US/09766399

: Patent No. US20010047092A1

: GENERAL INFORMATION:

: APPLICANT: Bruce, Wesley B.

: TITLE OF INVENTION: No. US20010047092A1 Plant Promoters and Methods of Use

: FILE REFERENCE: 1165

: CURRENT APPLICATION NUMBER: US/09/766,399

: PRIOR APPLICATION REMOVED - See file wrapper or Palm

: NUMBER OF SEQ ID NOS: 72

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 42

: LENGTH: 30

: TYPE: DNA

: ORGANISM: Zea mays

US-09-766-399-42

Query Match 9.3%; Score 30; DB 9; Length 30;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 AATCTACTCTCAACCAACCCAGCTTTGATCTGCTTACTGATCAACCAAGCTTC 184

DB 1 AATCTACTCTCAACCAACCCAGCTTTGATCTGATCTGATCTGATCTGATCTGATCTG 30

RESULT 15

US-09-766-399-43

: Sequence 43, Application US/09766399

: Patent No. US20010047092A1

: GENERAL INFORMATION:

: APPLICANT: Ruben et al.

: TITLE OF INVENTION: 62 Human Secreted Proteins

: FILE REFERENCE: P2039P1

: CURRENT APPLICATION NUMBER: US/10/050,704

: PRIOR APPLICATION REMOVED - See file wrapper or Palm

: NUMBER OF SEQ ID NOS: 344

: SOFTWARE: PatentIn Ver. 2.0

: SEQ ID NO 57

: TYPE: DNA

: ORGANISM: Homo sapiens

US-10-050-704-57

Query Match 9.3%; Score 29.8; DB 14; Length 1769;

Best Local Similarity 51.1%; Pred. No. 13;

Matches 70; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 152 TGCAATCTACTCTCAACCAACCCAGCTTTGATCTGCTTACTGATCAACCAAGCTTC 211

DB 1395 TGAACCT 1336

QY 212 CCTGATACGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 271

DB 1335 GCAGCTACGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276

QY 272 AAGAGAGAGGTTGGTCA 288

DB 1275 TGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1259

Query Match	100.0%	Score 322	DB 13	length 322
Best Local Similarity	100.0%	Prod. 1,4e+101		
Matches 322;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	GAATTCGTGCTGCGGGGTCAACTGAAACGGTAAACAGATGAAAGTGGATATCTCTTCTCT	60	
Db	1	GAATTCGTGCTGCGGGGTCAACTGAAACGGTAAACAGTGCAGAAACGGATCTCTTCTCT	60	
QY	61	CTCTGCAATCCGTCGCTGTGAGCAAAATGGGGCAATGCGCTACCTATACACCAACATAT	120	
Db	61	CTCTGCAATCCGTCGCTGTGAGAAATGAATGCGCTATCTATATATATATATATATAT	120	
QY	121	CACCTAGAAAGAGAGACGGCTCTCTGATATGATATGCAATATACCTCCCAACCAACCGACCTT	180	
Db	121	CACCTAGAAAGAGAGACGGCTCTCTGATATGATATGCAATATACCTCCCAACCAACCGACCTT	180	
QY	181	TGATATGCTTACATGATGATCAACAAATGTGTGATATCATGTGATATATTTT	240	

Patent No. 635919

Patent No. 635919

```

1 COMPITE-READABLE FORM:
2
3 MEDIUM TYPE: Floppy disk
4
5 COMPUTER: IBM PC compatible
6
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8
9 SOFTWARE: Patent Release #1.0, Version #1.30
10
11 CURRENT APPLICATION DATA:
12
13 APPLICATION NUMBER: US/08/413,135
14
15 FILING DATE:
16
17 CLASSIFICATION: 800
18
19 ATTORNEY/AGENT INFORMATION:
20
21 NAME: Seay, Nicholas J
22
23 REGISTRATION NUMBER: 27,386
24
25 REFERENCE/DOCKET NUMBER: 960296.92808
26
27 TELECOMMUNICATION INFORMATION:
28
29 TELEPHONE: 608-251-5000
30
31 TELEFAX: 608-251-9166
32
33 INFORMATION FOR SEQ ID NO: 1.
34
35 SEQUENCE CHARACTERISTICS:
36
37 LENGTH: 3183 base pairs
38
39 TYPE: nucleic acid
40
41 STRANDEDNESS: double
42
43 TOPOLOGY: linear
44
45 MOLECULE TYPE: other nucleic acid
46
47 DESCRIPTION: /desc = "SAC12-1 Promoter DNA"
48
49 US-08-413-135-1

```

Query Match	9.98,	Score 32,	DB 1,	Length 3183,
Best Local Similarity	50.7%;	Pred. No. 0.44;		
Matches 77;	Conservative	0;	Mismatches 75,	Indels 0;
			Gaps	0

QY	12	TCACCTAGAAAAACGCGCGCTCTCGATTCGATTCGAAATTCACCCCAACCAACCCAGCT	179
Db	424	TCATATATATACATCATTCATGATCATATTTAGATATGATCTACTAATTCATATACATCC	365
QY	180	TTTATCTCTTACTGTATATCCCAAGTTGGTCTGATACCATGTTGAGATATATGCTCTT	239
Db	364	ATTCAACCTGTTTCATTATAAAAACAGATTCGTATATAAACCTGGAAATATTTGTTTTT	305
QY	240	TCTTCTAGAAATGTTCTCCCGCATGCTTAA	271
Db	304	AATTATATTGAAACGTACATCTTTATGGGCTCT	273

RESULT 4
 US-08-987-367-3/C
 Sequence 3, Application US/08987367
 Patent No. 6399859
 GENERAL INFORMATION:
 APPLICANT: Nichols, Scott
 APPLICANT: Dhugra, Kanarpal S.
 APPLICANT: Singletary, George William
 APPLICANT: Saunders, Court
 APPLICANT: Wilcher, Derrick
 APPLICANT: Bruce, Wesley B.
 APPLICANT: Sims, Lynne
 APPLICANT: Lu, Guihua
 APPLICANT: Zhong, Gan-Yuan
 TITLE OF INVENTION: PLANT URIDINE DIPHOSPHATE-GLUCOSE
 NUMBER OF INVENTION: DEHYDROGENASE GENES, PROTEINS, AND USES THEREOF
 NUMBER OF SPOUNCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

19 TCACTGACCCGTAACAGTGGAAAGTATCTTCTCTCTGCAATCCGTCGCT 78
1401923 TTACTTAATTCCTTAACACATCCGGAAGGAAATATTACTGCTACCAAGACCTACTT 1401864
79 GGAACAAATATGCTGAGTCTCTCTATATATACACCAATATATACATGAAAGGAG 137
1401863 GCAAGAAATGAAAGCTAAAGTAAATATATGAAATCCCAAAATATATGTAAGG 1401805

RESULT 10
US-09-328-352-3148/C
Sequence 3148, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3148
LENGTH: 699
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3148

Query Match 9.1%; Score 29.2; DB 4; Length 699;
Best Local Similarity 56.1%; Pred. No. 2.1;
Matches 55; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

224 TCGCATTTATGCTCTCTCTCTAGCAATGTCCTCGGATCGTTTAAAGAGAGCTG 283
576 TACATTAATTAAGAGCTTTTTCATCGCTGATTATGAACAGCTTTCAATAAGATGTCG 517
284 GTCAGCATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
516 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479

RESULT 11
US-09-197-063-1/C
Sequence 1, Application US/09197063
Patent No. 6261817
GENERAL INFORMATION:
APPLICANT: Zalacatin, Magdalena
APPLICANT: Brown, James R.
APPLICANT: Biswas, Sanjoy
APPLICANT: Warren, Richard L.
APPLICANT: Shilling, Lisa K.
TITLE OF INVENTION: No. 6261817el GuAA
FILE REFERENCE: GM10121
CURRENT APPLICATION NUMBER: US/09/197,063
CURRENT FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/066,350
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows version 3.0
SEQ ID NO 1
LENGTH: 1551
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION (1)... (1548)
US-09-197-063-1

Query Match 9.1%; Score 29.2; DB 3; Length 1551;
Best Local Similarity 52.5%; Pred. No. 3;
Matches 64; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
61 CTCTGCAATCCGTCGCTGGAAGCAAAATGGCGAGTGGCTACTTATACACCAACTTAT 120

1330 CAGTGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1271
121 CACCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
1270 CAAAGAAATATAGTCTGATTAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211
181 TG 182
1210 TG 1209

RESULT 12
US-09-221-294-3
Sequence 3, Application US/09221294
Patent No. 6268138
GENERAL INFORMATION:
APPLICANT: Riccardo Dalla-Favera and
Alessandro Massimo Gianni
TITLE OF INVENTION: A Retroviral Vector Capable of Transducing the
TITLE OF INVENTION: Aldehyde Dehydrogenase-1 Gene and Uses of Said
TITLE OF INVENTION: Vector
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,294
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42990-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2904 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: N
ANTI-SENSE: N
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 363..2274
OTHER INFORMATION:
US-09-221-294-3

Query Match 9.1%; Score 29.2; DB 3; Length 2904;
Best Local Similarity 53.5%; Pred. No. 4;
Matches 61; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

26 AACGCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 85
844 AACCCACCCAGTGTGAAGGAGAGCTTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 903
86 AATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 139
904 ACAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 957

RESULT 13

US-09-302-620B-94
Sequence 54, Application US/09302620B
Patent No. 6331420

GENERAL INFORMATION:

APPLICANT: Wilson, C. Ron
APPLICANT: Craft, David L.
APPLICANT: Elrich, Dudley
APPLICANT: Espoo, Mark
APPLICANT: Madduri, Krishna M.
APPLICANT: Cornett, Cathy A.
APPLICANT: Brenner, Alfred A.
APPLICANT: Tang, Maria
APPLICANT: Loper, John C.
APPLICANT: Gleeson, Martin

TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS
TITLE OF INVENTION: RELATING THERETO

FILE REFERENCE: 1010-16.seq
CURRENT APPLICATION NUMBER: US/09/302.620B
CURRENT FILING DATE: 1999-04-30

NUMBER OF SEQ ID NOS: 109
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 94

LENGTH: 3348

TYPE: DNA

ORGANISM: Candida tropicalis
US-09-302-620B-94

Query Match
Best Local Similarity 59.8%; Pred. No. 4.3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 241 CTCTCTAGACATGCTGCGCATGCTTATTAAGACAGCTGCGACGATGATCTG 300
DB 76 CTGCTTCAACACCTTATCTGCTATTTGTCATGCAACGCTGCTGCAACGATATCTG 135
QY 301 CCAAGTATATACGACGACGCTG 322
DB 136 CCAAGTATATACGACGACGCTG 157

RESULT 14

US-09-912-161-12
Sequence 12, Application US/09912161
Patent No. 6503734

GENERAL INFORMATION:

APPLICANT: David, Craft L.
APPLICANT: Madduri, Krishna
APPLICANT: Loper, John C.
TITLE OF INVENTION: CYTOCHROME b5 GENE AND PROTEIN OF CANDIDA TROPICALIS AND METHODS
TITLE OF INVENTION: THEREFO
FILE REFERENCE: M6368 (1010-35)
CURRENT APPLICATION NUMBER: US/09/912.161
CURRENT FILING DATE: 2001-07-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentin version 3.1
SEQ ID NO 12
LENGTH: 3348
TYPE: DNA
ORGANISM: Candida tropicalis
US-09-912-161-12

Query Match
Best Local Similarity 9.1%; Score 29.2; DB 4; length 3348;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 241 CTCTCTAGACATGCTGCGCATGCTTATTAAGACAGCTGCGACGATGATCTG 300
DB 76 CTGCTTCAACACCTTATCTGCTATTTGTCATGCAACGCTGCTGCAACGATATCTG 135

QY 301 CCAAGTATATACGACGACGCTG 322
DB 136 CCAAGTATATACGACGACGCTG 157

RESULT 15

US-08-961-527-159/c
Sequence 159, Application US/08961527
Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527

FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PR340P1
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO. 159.

SEQUENCE CHARACTERISTICS:
LENGTH: 9048 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-159

Query Match
Best Local Similarity 9.1%; Score 29.2; DB 4; length 9048;
Matches 64; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 61 CTCTGCAATCGTGCGGTGGAAGCAATGCGCGAGTCGCTACTATATGACGCAACTTAT 120
DB 8077 CTGCTTCAACGAGTGAAGTATTCGCAATATTCGCGGTCAAGTCACACTTACGATTTCTT 8018
QY 121 CACTTAAAGAGCGACGCGCTGCTGATTCATTTGCAATTCATTCGCAACGCAACGCTT 180
DB 8017 CACGAAGATAGGTGATTCAGCAATGATTTTAAATTTTCTTCAGTGAATTCAGCTGA 7958
QY 181 TG 182
DB 7957 TG 7956

Search completed: August 18, 2003, 14:30:51
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 18, 2003, 13:14:55 ; Search time 1913 Seconds
(without alignments)
4090.978 Million cell updates/sec

Title: US-10-058-566-5

Perfect score: 322

Sequence: 1 gaattcgtctgcgcgctc.....agtgtctagctgagaacatg 322

Scoring table: IDENTITY_NTC Gapop 10 0 Gapext 1 0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

EST:
1: em_estba:
2: em_esthum:
3: em_estln:
4: em_estmu:
5: em_estov:
6: em_estopl:
7: em_estro:
8: em_hic:
9: qb_est1:
10: qb_est2:
11: qb_hic:
12: qb_est3:
13: qb_est4:
14: qb_est5:
15: em_estom:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pin:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_tod:
26: em_gss_ping:
27: em_gss_vrl:
28: qb_gss1:
29: qb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	223.8	69.5	740	29 CC349513 OGOAS63TH
2	190.8	59.3	508	29 B2401831 B2401831
3	190.8	59.3	524	29 B2401828 B2401828
4	45.8	14.2	685	12 B211851 B211851

C 5	38.2	11.9	533	9	AW784857	AW784857 115155 MA
C 6	37	11.5	849	13	B0745750	B0745750 CH2#002.A
C 7	36.8	11.4	664	10	B8449721	B8449721 B8449721
C 8	36.8	11.4	675	9	AV349139	AV349139 AV349139
C 9	36.8	11.4	2560	11	AK032641	AK032641 Mus muscu
C 10	36.8	11.4	2574	11	AK051108	AK051108 Mus muscu
C 11	36	11.2	1053	13	B0745776	B0745776 CH2#002.C
C 12	36	11.2	1069	13	B0745749	B0745749 CH2#002.A
C 13	35.6	11.1	549	28	AA793738	AA793738 ZM0047K08
C 14	35.4	11.0	607	12	B0543444	B0543444 B0543444
C 15	35.4	11.0	612	12	B0101888	B0101888 B0101888
C 16	35.4	11.0	627	12	B0101944	B0101944 B0101944
C 17	35.2	10.9	588	10	BC524736	BC524736 6-60 Stev
C 18	34.8	10.8	935	29	CNS03A3N	AL136588 Drosophi1
C 19	34.6	10.7	522	12	B7501294	B7501294 B7501294
C 20	34.4	10.7	284	10	BH245583	BH245583 B8246583
C 21	34.4	10.7	665	12	B0529082	B0529082 B0529082
C 22	34.4	10.7	803	13	B0745777	B0745777 CH2#002.C
C 23	34	10.6	385	12	BH798289	BH798289 K-EST0081
C 24	33.8	10.5	859	28	BH655718	BH655718 HOHVH587R
C 25	33.6	10.4	265	9	AA348378	AA348378 EST55183
C 26	33.6	10.4	671	28	BH453718	BH453718 B060M19TF
C 27	33.6	10.4	800	28	BH488995	BH488995 B0H8236TF
C 28	33.6	10.4	560	29	CNS080PA	AL053263 Drosophi1
C 29	33.4	10.4	608	12	B026852	B026852 B026852
C 30	33.4	10.4	1201	29	CNS016CY	AL196588 Drosophi1
C 31	33.2	10.3	348	14	T01667	T01667 WEST02388.E
C 32	33.2	10.3	482	9	AM653990	AM653990 102949 MA
C 33	33.2	10.3	751	28	BH700981	BH700981 B0MGR44TF
C 34	33.2	10.3	685	28	BH461211	BH461211 B0MPS29TF
C 35	33	10.2	408	14	118714	118714 5C04G11-F21
C 36	33	10.2	537	28	A0765571	A0765571 HS-5337.A
C 37	33	10.2	540	14	CB144345	CB144345 K-EST0198
C 38	33	10.2	1192	12	BG787172	BG787172 SEADMC007
C 39	33	10.2	1201	9	AL552063	AL552063 AL552063
C 40	32.8	10.2	453	9	AA941154	AA941154 LD25138.5
C 41	32.8	10.2	632	29	B2905637	B2905637 CH240-27F
C 42	32.8	10.2	724	28	BH695201	BH695201 B0MET80TR
C 43	32.8	10.2	742	29	AG085810	AG085810 Pan. trogl
C 44	32.8	10.2	768	28	BH547822	BH547822 B010085TF
C 45	32.8	10.2	899	29	CNS034200	AL225105 Tetradon

ALIGNMENTS

RESULT 1
LOCUS CC349513 740 bp DNA linear GSS 16-MAY-2003
DEFINITION OGOAS63TH_ZM_0.7.1.5_KB Zee mays genomic clone ZMBMB0368105,
genomic survey sequence.
ACCESSION CC349513
VERSION CC349513.1 GI:30818920
KEYWORDS GSS.
SOURCE Zee mays
ORGANISM Zee mays
Eukaryota; Viridiplantae; Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta; Magnoliophyta; Liliopsida, Poales, Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
,A., Fraser,C.M., Budiman,M.A., Rodell,J.A., Rohlfing,T., Citek
,P.W., Nunberg,A., Pobbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished
COMMENT Contact: Cathy Whitelaw
TIOR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@ltd.org
Seq primer: TR
Class. sheared ends.

```
FEATURES
  source
    location/Qualifiers
    1..740
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone_lib="ZM_0.7-1.5_KB"
    /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

BASE COUNT
  187 a 184 c 164 g 205 t

Query Match
  Best Local Similarity 85.6%, Pred. No. 8.5e 59;
  Matches 273; Conservative 0; Mismatches 42; Indels 4; Gaps 2;

OY 6 CTGCTTCGGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTCTG 65
    |||||
DB 1 CTGCTTCGGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTGCAAT 60
    |||||
OY 66 CAATCCGTCGCGTGAAGCAATATGCGCATCTATATACACCACTTATACCT 125
    |||||
DB 61 CCGT--AAAGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTGCAAT 118
    |||||
OY 126 AGAAAGCGAGCGGTCTGTGATGATTCGAATCTACCTCAACCAACCCAGCTTTGAT 185
    |||||
DB 129 ACAAAGCGAGCGGTCTGTGATGATTCGAATCTACCTCAACCAACCCAGCTTTGAT 178
    |||||
OY 186 CTGCTTACGTATACCAACAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 245
    |||||
DB 179 CTGCTTACGTATACCAACAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 238
    |||||
OY 246 CTAGATGTTCTGCGCATGCTTATATAGAGAGTGTGTACG--ATCGATCTGCGCA 303
    |||||
DB 219 CTAGATGTTCTGCGCATGCTTATATAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 298
    |||||
OY 304 CTGCTTACGTATACCAACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 322
    |||||
DB 299 GTGCTAGTGTAGAGACATG 317

RESULT 2
LOCUS BZ401831 508 bp DNA linear GSS 04-DEC-2002
DEFINITION OGAAG51TM ZM_0.7-1.5_KB Zea mays genomic clone ZMMBMA0015G20,
ACCESSION BZ401831
VERSION BZ401831.1 GI:26026901
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE
  Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 508)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
  ,R.W., Nunberg,A., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
  Consortium for Maize Genomics
  Unpublished
  Other GSSs: OGAAG51TC
  TIGR
  Contact: Cathy Whitelaw
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
    1..508
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone_lib="ZM_0.7-1.5_KB"
    /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

BASE COUNT
  135 a 129 c 107 g 137 t

Query Match
  Best Local Similarity 83.8%, Pred. No. 1.5e 48;
  Matches 228; Conservative 0; Mismatches 42; Indels 2; Gaps 1;

OY 6 CTGCTTCGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTCTG 65
    |||||
DB 239 CTGCTTCGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTGCAAT 298
    |||||
OY 66 CAATCCGTCGCGTGAAGCAATATGCGCATCTATATACACCACTTATACCT 125
    |||||
DB 299 CCGT--AAAGCGGTCAACCTGAACCCGTAACAGTCGCAAGTGTACTCTTCTCTGCAAT 356
    |||||
OY 126 AGAAAGCGAGCGGTCTGTGATGATTCGAATCTACCTCAACCAACCCAGCTTTGAT 185
    |||||
DB 357 ACAAAGCGAGCGGTCTGTGATGATTCGAATCTACCTCAACCAACCCAGCTTTGAT 416
    |||||
OY 186 CTGCTTACGTATACCAACAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 245
    |||||
DB 417 CTGCTTACGTATACCAACAGTTGTGATGATGATGATGATGATGATGATGATGATGAT 476
    |||||
OY 246 CTAGATGTTCTGCGCATGCTTATATAGAGA 277
    |||||
DB 477 CTAGATGTTCTGCGCATGCTTATATAGAGA 508

RESULT 3
LOCUS BZ401828/c 524 bp DNA linear GSS 04-DEC-2002
DEFINITION OGAAG51TC ZM_0.7-1.5_KB Zea mays genomic clone ZMMBMA0015G20,
ACCESSION BZ401828
VERSION BZ401828.1 GI:26026898
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE
  Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 524)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick
  ,R.W., Nunberg,A., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
  Consortium for Maize Genomics
  Unpublished
  Other GSSs: OGAAG51TM
  TIGR
  Contact: Cathy Whitelaw
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TP
  Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
    1..524
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone_lib="ZM_0.7-1.5_KB"
    /note="Vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"
```


REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 664)	Mammalia, Eutheria, Rodentia, Sciurognathi; Muridae, Murinae, Mus.			
	Atakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M., and Hayashizaki,Y.	RIKEN Mouse hSTS (Atakawa,T., et al. 2001)		
		Unpublished		
		On Jul 21, 2000 this sequence version replaced gi:9310756		
		Contract: Yoshinide Hayashizaki		
		Laboratory for Genome Exploration Research Group, RIKEN Genomic		
		Sciences Center(GSC), Yokohama Institute		
		The Institute of Physical and Chemical Research (RIKEN)		
		1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan		
		Tel: 81-45-503-9222		
		Fax: 81-45-503-9216		
		Email: genome-res@gsc.riken.go.jp,		
		URL: http://genome-jsc.riken.go.jp/		
		Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh		
		,M., Komno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.		
		Normalization and subtraction of cap-trapper-selected cDNAs to		
		prepare full-length cDNA libraries for rapid discovery of new		
		genes. Genome Res. 10 (10), 1617-1630 (2000)		
		wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,F.,		
		Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura		
		S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and		
		Hayashizaki,Y.		
		RIKEN integrated sequence analysis (RISA) system--384-format		
		sequencing pipeline with 384 multicapillary sequencer. Genome Res.		
		10 (11), 1757-1771 (2000)		
		Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara		
		,Y., and Hayashizaki,Y.		
		Computer-based methods for the mouse full-length cDNA		
		encyclopedia: real-time sequence clustering for construction of a		
		nondundant cDNA library. Genome Res. 11 (2), 281-289 (2001)		
		Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanka,I., Aizawa		
		,K., Fukuda,S., Hara,A., Itoh,M., Kawai,T., Shibata,K., and		
		Hayashizaki,Y.		
		Computational Analysis of Full-Length Mouse cDNAs Compared with		
		Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)		
		Please visit our web site (http://genome-jsc.riken.go.jp/) for		
		further details.		
		cDNA library was prepared and sequenced in Mouse Genome		
		Encyclopedia Project of Genome Exploration Research Group in Riken		
		Genomic Sciences Center and Genome Science Laboratory in Riken		
		Division of Experimental Animal Research in Riken contributed to		
		prepare mouse tissues.		
		Location/Qualifiers		
		1. 664		
		/organism="Mus musculus"		
		/mol_type="mRNA"		
		/db_xref="taxon:10090"		
		/clone="D030073E24"		
		/dev_stage="9 days embryo"		
		/lab_host="DH10B"		
		/clone_lib="RIKEN full-length enriched, 9 days embryo"		
		/note="Site.1: SalI; Site.2: BamHI; cDNA library was		
		prepared and sequenced in Mouse genome Encyclopedia		
		Project of Genome Exploration Research Group in Riken		
		Genomic Sciences Center and Genome Science Laboratory in		
		RIKEN. Division of Experimental Animal Research in Riken		
		contributed to prepare mouse tissues. 1st strand cDNA was		
		primed with a primer [5'		
		GAGGAGAGAGCGCGCCACATCGATTTTTTTTTTTTTTTTNN 3'], cDNA was		
		prepared by using trehalase thermo activated reverse		
		transcriptase and subsequently enriched for full-length by		
		cap-trapper. Second strand cDNA was prepared with the		
		primer adapter of sequence [5'		
		GAGGAGAGATTCGAGTTAATTAATTCGCCCCCCCCCCC 3']. cDNA		

1 1053

BASE COUNT
ORIGIN

Best Local Similarity	67.18;	Pred No. 5.2;			
Matches	51; Conservative	0; Mismatches	25; Indels	0; Gaps	0

Db 690 TCCCAAGTTGAGCTGAGAGCTGGAAGACTTAGATTGTCTTTGCCATTAATGTTCT 749

Db 750 GGGCCTGGTTAGAAG 765

DEFINITION	ACCESSION
CH2#002_A04.3 canine heteronormalized cDNA library in pRiScript	U0755740
Canis familiaris cDNA clone CH2#002_A04.3', mRNA sequence.	U0755740

KEYWORDS	EST.
SOURCE	Canis familiaris (dog)

REFERENCE
Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 1069)

JOURNAL Unpublished
COMMENT Other_ESTS: CH2#002_A04T7

Tel: 615 936 2660
Fax: 615 936 2661

High quality sequence start: 59
High quality sequence stop: 826.

```
FEATURES
source      location
1. .1065
```

BASE COUNT	255 a	286 c	234 g	294 t
ORIGIN				

Best Local Similarity 67.1%; Pred. No. 5.2;
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Db 656 TCCCAAGATTGAGCTGACGAGCTGGACACTTAGATTGTCTTTGCCAATACTTTCC 715

Db 716 GGGCCTGCTTAGAAC 731

DEFINITION ZM0047K08F Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG2M0047K08 F, genomic survey sequence.

KEYWORDS	GSS.
SOURCE	Mus musculus (house mouse)

REFERENCE
1 (bases 1 to 549)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.

JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Class: plasmid ends
High quality sequence stop: 549.

```
/clone="UUGC2M0047K08"  
/sex="Male"
```

/lab host="E. coli strain XL10-gold, T1-resistant, F-"
 /clone.lib="Mouse 10kb plasmid UUCGM library"
 /note="Vector: pMD42nv; purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114) [pAR12972.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT	172 a	120 g	166 t
ORIGIN	91 c	120 g	166 t

Query Match 11.1%; Score 35.6; DB 28; Length 549;
 Best Local Similarity 62.2%; Pred. No. 5;
 Matches 56; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY	208	TTCTGCTGATACGATGTCGCATTAATGCTGTTCTGTCAGACATGTCCTCCGACACT	267
Db	281	TTTCTTTCGATGATATTCTTCTTATTTTCTTTTAAATTTTATACAAATTTTTCACAACT	222

OY	268	TTATAGACAGAGCTTGCTACCATGATCT	297
Db	221	TTACAGCCCTCCCGACGTCATCAGCACT	192

RESULT 14
 LOCUS HJ343444
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oryzias latipes (Japanese medaka)
 Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Altherinomorpha; Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
 1 (bases 1 to 607)
 Kohara,Y., Shin-I,T., Kimura,T., Marita,T., Jinbo,T. and Takeda,H.
 Medaka EST Project in Takeda's lab
 Unpublished
 Contact: Tadasu Shin-I
 Center for Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel.: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genetics.nig.ac.jp.
 Location/Qualifiers
 1..607
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="Hd-ir"
 /db_xref="taxon:8090"
 /clone="MF01SSH030G19"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="segmentation stage 20 - 25"
 /clone.lib="MF01SSB cDNA"

```

BASE COUNT      132 a     164 c     183 g     127 t     1 others
ORIGIN
Query Match          11.0%, Score 35.4; DB 12; Length 607,
Best Local Similarity 59.4%; Pred No 6.1;
Matches 60; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

UY 188 GTTATTGATCAACAAAGTTCTTCATATCATGTGATATATGCTTTCTCTCT 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GTTAATGATGACAAGAAGGTTGGACACACACAAACAGCCCTCCTACTCTCTCTCT 60

UY 248 ACAAATCTCTCCAGATCGCTTATAAGCAAGAGTTGGTAG 288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GATATGTTCTGCATATGATGATTAATTTAAGAGAGTGCGAG 101

RESULT 15
BU012888/c        612 bp   mRNA       linear   EST 05-DEC-2001
DEFINITION        BU012888 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA181M1 5',
LOCUS              BU012888
ACCESSION          BU012888
VERSION            BU012888.1 GI:17363321
KEYWORDS           EST.
SOURCE             Oryzias latipes (Japanese medaka)
ORGANISM           Oryzias latipes
COMMENT            Contact: Tadasu Shin-i
                    Center For Genetic Resource Information
                    National Institute of Genetics
                    1111 Yata, Mishima, Shizuoka 411-8540, Japan
                    Tel.: 81-559-81-6856
                    Fax: 81-559-81-6855
                    Email: tsunigenes.nig.ac.jp.
                    Location/Qualifiers
                        1..612
                        /organism="Oryzias latipes"
                        /mol_type="mRNA"
                        /strain="Md-ir"
                        /db_xref="taxon:8090"
                        /clone="MF01SSA181M1"
                        /sex="mixture of female and male"
                        /tissue.type="whole embryo"
                        /dev_stage="segmentation stage 20 - 25"
                        /clone_idb="MF01SSA cDNA"

FEATURES
Source
1..612
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Md-ir"
/db_xref="taxon:8090"
/clone="MF01SSA181M1"
/sex="mixture of female and male"
/tissue.type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_idb="MF01SSA cDNA"

Query Match          11.0%; Score 35.4; DB 12; Length 612;
Best Local Similarity 59.4%; Pred No. 6.1;
Matches 60; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

UY 188 GTTATTGATCAACAAAGTTCTTCATATCATGTGATATATGCTTTCTCTCT 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 602 GTTAATGATGACAAGAAGGTTGGACACAAATATGTCCTCTACTCTCTCTCTCT 543

UY 248 ACAATGTTCTGCCATGCTTATATACGAAGAGTTGCTGACG 288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 542 GAATGTGTTCTGATATGCAATATTTAAGAGAGTCTAGAG 502

Search completed: August 18, 2003, 14:29:32
Job time : 1918 secs
```

Now I got to know for what we were doing.

New isolated DNA molecule, useful for mediating male fertility in maize plants to produce a hybrid that may increase the yield of desirable

The invention relates to a DNA molecule mediating male fertility in CC plants comprising the HS92-7 gene, represented by American Type Culture Collection (ATCC) deposit number 98932. The DNA molecules are useful in mediating male fertility in plants, particularly maize, to produce a hybrid plant that may increase the yield and combination of desirable characteristics in plants, such as resistance to disease and insects or heat and drought tolerance. The present sequence represents the maize HS92-7 promoter sequence.

Query Match	100.0%	Score 322	DB 24	Length 322
Host Local Similarity	100.0%	Prod. No. 4.4e-101		
Matches 322	Conservative 0	Mismatches 0	Indels 0	Gaps 0

DY 1 GAATCTGTGTTGGAGGTAACTGAACCGTAAACGTGGAAAGTGATACTCTTTCT 60
|||||
|||
Dd 1 GAATTCTGCTCTGGCGGTAACTGAACCGTAAACGTGGAAAGTGATACTCTTTCT 60

Qy 61 CTCTGCAATTCCTGTCCTGCAAGCAATGGCGCAGTCGCGCTTATTCACACCAACTTAT 120
|||||
Dd 61 CTCTGCATCCGTGCGCTGCAAGCAATGGCGCAGTCGCGCTTATTCACACCAACTTAT 120

Qy 121 CACCTAGAAAAGCGACGGCTGCTGATGATTGCAATCTACTCCAAACCAACCCAGCTT 180
|||||
121 CACCTAGAAAAGCGACGGCTGCTGATGATTGCAATCTACTCCAAACCAACCCAGCTT 180

QY 181 TGATCTGCTTACTGTGATCACCMAAGTGTGCTGATAGATGTGGGATTATGCTCTTT 2400
|||||
181 TGATCTGCTTACTGTGATCACCMAAGTGTGCTGATAGATGTGGGATTATGCTCTTT 2400

QY 241 CTTCCTAGAAATGTTCCGCCGATCCTTTATAGACAAGCTTGGTCAGCATTCGATCTCTG 300
|||||
Dδ 241 CTCTCTAGAAATGTTCCGCCGATCCTTTATAGACAAGCTTGGTCAGCATTCGATCTCTG 300

QY	301	CCAGTGTCTAGCTGAGAACATG	322
nb	301	CCAGTGTCTAGCTGAGAACATG	332

RESULT 2	
ABQ79567	
ID	ABQ79567 standard; DNA; 2541 BP.

AC	AHQ79567;
XX	
DT	25-NOV-2002 (first entry)

DN: % . mayas HS92-7 genomic DNA sequence.

KM HS92-7, fertility, plant; disease resistance; drought tolerance;
KM maize; gene; ds.

05 Zea mays.

PN WO200263021-A2
XX

15-AUG-2002.

30-JAN-2002; 2002W0-01502713

PR 08-FEB-2001; 200105-26/521P
XX

XX (PION-) PIONEER HI-BRED INI INC.

[illegible]

DR WP1; 2002-657541/70.

DR P-PSDB; ABB81980

PT New isolated DNA molecule, useful for mediating male fertility in maize
PT plants to produce a hybrid that may increase the yield of desirable
PT characteristics in plants, e.g resistance to disease and insects or
PT heat and drought tolerance -

PS Claim 3; Fig 5A-B; 54pp; English.

CC The invention relates to a DNA molecule mediating male fertility in
CC plants comprising the BS92-7 gene, represented by American Type Culture
CC Collection (ATCC) deposit number 98932. The DNA molecules are useful in
CC mediating male fertility in plants, particularly maize, to produce a
CC hybrid plant that may increase the yield and combination of desirable
CC characteristics in plants, such as resistance to disease and insects or
CC heat and drought tolerance. The present sequence represents the maize
CC BS92-7 genomic DNA sequence.

Sequence 2541 RP; 577 A; 693 C; 631 G; 640 T; 0 others;

Query Match:	100.0%;	Score 322;	DB 24;	Length 2541;
Best Local Similarity	100.0%;	Pred No 1	le-100;	
Matches 322;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

OY 1 GAATTCCTCTCTGGGGGTCACGGAACCCGTTAAACAGTGGAAAGTCGATTACTCTTTTCCT 60
 |||||
 1 GAATTCCTCTCTGGGGGTCACGGAACCCGTTAAACAGTGGAAAGTCGATTACTCTTTTCCT 60
 Db 1 GAATTCCTCTCTGGGGGTCACGGAACCCGTTAAACAGTGGAAAGTCGATTACTCTTTTCCT 60

Gy 61 CTCTGCAATTCGTGGTCCGTGGAAACAAATCGCGCATGTTACTATTATCACAGCAACTTAT 120
|||||
Db 61 CTCGCAATCCGTGCCCGTGGAAGCAAATGGCGGCAGCTCGCCTACTTATACACCAACTTAT 120

Qy 121 CACCTAGAAAAGCAGCGCTTCCTGGATTCGATTGCAATCTACCTCCACCACCCAGCTT 180
|||||
Db 121 CACCTAGAAAAGCAGCGCTTCCTGGATTCGATTGCAATCTACCTCCACCACCCAGCTT 180
|||||

Qy 181 TGATCTGCTACTGTGATCACAAGTGTGCGCATACAGATGTGGCATTAATGCTCTTT 240
|||||
Db 181 TGATCTGCTACTGTGATCACAAGTGTGCGCATACAGATGTGGCATTAATGCTCTTT 240

QY 241 CTCTCTAGATGTCTGCGAGTTTATAAGAGAACTTGGTCACATGATCTCTG 300
 |||||
 Eb 241 CTCTCTGAATGTCTCCGCGAGCTTTATTAAGAGAACTTGGTCACATGATCTCTG 300
 |||||

QY	301	CCAGTGTCTAGCTGAGACATG	322
Pb	301	CCAGTGTCTAGCTGAGACATG	322

RESULT 3	
ABQ79569	
ID	ABQ79569 standard; DNA; 187 BP.

AC ABQ79569

25-NOV-2002 (first entry)

2. *mays* BS92-7 promoter essential region sequence

KW BS92-7, fertility; plant; disease resistance; drought tolerance, maize; promoter; ds

XX 7.22 mays

XX
PN WC200263021-A2

AA 15-AUG-2002.
PD

AA 30-JAN-2002; 2002WO-US02713.
PF

PR 08-FEB-2001; 2001US-267527P.

PA (PION-) PIONEER HI-BRED INT' INC.

XX Albertsen M, Fox T, Hufman G, Trimmell M;
 XX
 XX WPI: 2002-657541/70.
 DR
 PT New isolated DNA molecule, useful for mediating male fertility in maize
 PT plants to produce a hybrid that may increase the yield of desirable
 PT characteristics in plants, e.g. resistance to disease and insects or
 PT heat and drought tolerance.
 XX
 PS Claim 38: Fig 10; 54pp; English.
 XX
 CC The invention relates to a DNA molecule mediating male fertility in
 CC plants comprising the BS92-7 gene, represented by American Type Culture
 CC Collection (ATCC) deposit number 98932. The DNA molecules are useful in
 CC mediating male fertility in plants, particularly maize, to produce a
 CC hybrid plant that may increase the yield and combination of desirable
 CC characteristics in plants, such as resistance to disease and insects or
 CC heat and drought tolerance. The present sequence represents the maize
 CC BS92-7 promoter essential region sequence.
 CC
 SO Sequence 187 BP; 42 A; 46 C; 40 G; 59 T; 0 other;

Query Match 58.1%; Score 187; DB 24; Length 187;
 Best Local Similarity 100.0%; Pred. No. 1.9e-54;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 CCGGCTCGATGATGCTCAATCTACCTCCACCAACCCAGCTTTGTATCTGCTTACG 195
 DB 1 CCGGCTCGATGATGCTCAATCTACCTCCACCAACCCAGCTTTGTATCTGCTTACG 60
 QY 196 TGATCAGCAAGCTTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 255
 DB 61 TGATCAGCAAGCTTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 256 CCGGCGATGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315
 DB 121 CCGGCGATGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 316 GAACATG 322
 DB 181 GAACATG 187

RESULT 4
 ABL10980/C
 ID ABL10980 standard; cDNA; 3413 BP.
 XX
 AC ABL10980;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 27422.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PH CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 PT
 DR WPI: 2001-656860/75.

DR P-PSDB; ABB66877.
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Claim 1: SEQ ID NO 27422; 41pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pcl_sequences.
 CC
 SO Sequence 3413 BP; 997 A; 759 C; 673 G; 984 T; 0 other;

Query Match 11.1%; Score 35.6; DB 23; Length 3413;
 Best Local Similarity 51.2%; Pred. No. 0.19;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 154 CAATCTACCTCCACCAACCCAGCTTTGTATCTGCTTACCTGATCAGCAACTGCTGC 213
 DB 402 CATTTCTACTTAAAGCAGCAGCAAGTAATATCCCTCATGCTCTCATTCAGCAAGTAA 343
 QY 214 TGATCAGCTTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 273
 DB 342 TCACGGAAGTAATGCAAGATGAGATGATGATGATGATGATGATGATGATGAT 283
 QY 274 GAGCAAGCTTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 315
 DB 282 GGTATGCTTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 241

RESULT 5
 ABL12129/C
 ID ABL12129 standard; cDNA; 4329 BP.
 XX
 AC ABL12129;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 30869.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PH CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 PT
 DR WPI: 2001-656860/75.
 DR P-PSDB; ABB68026.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX CC Claim 1; SEQ ID NO 30866; 21bp + Sequence Listing; English.
 PS CC The invention relates to an isolated nucleic acid detection reagent
 XX CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB857737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX SQ Sequence 4329 BP; 1231 A; 1000 C; 1027 G; 1071 T; 0 other;
 Query Match 11.0%; Score 35.4; DB 23; Length 4329;
 Best Local Similarity 57.8%; Pred. No. 0.25;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 OY 43 AGTCGATACCTCTTCTCTCTGCAATCCGCGCGGCAAGCAAGATCGCGCAGTCCCTA 102
 Db 2644 ATTGGAACCTTTCTTCTCTCGAGCAATCCGAGCTTGTTCAGCAATGACGACACGCCGTT 2585
 OY 103 CTTATACACCAACTTATACCTAGAAAAGCGCGCTCGATGAT 151
 Db 2584 CGTCGCCAACAACCTTAATCTTAATATACGACGATCGCTCGATTAAT 2536
 RESULT 6
 AB12128
 ID AB1.12128 standard; cDNA; 9404 BP.
 AC AB1.12128;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 30866.
 XX
 KW *Drosophila*, developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS *Drosophila melanogaster*.
 XX
 FN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 DR WPI; 2001-656860/75.
 DR P-PSDB: ABB68025.
 XX
 PT New, isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 30866; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB857737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX SQ Sequence 9404 BP; 2675 A; 1973 C; 1984 G; 2772 T; 0 other;
 Query Match 11.0%; Score 35.4; DB 23; Length 9404;
 Best Local Similarity 57.8%; Pred. No. 0.36;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 OY 43 AGTCGATACCTCTTCTCTCTGCAATCCGCGCGGCAAGCAAGATCGCGCAGTCCCTA 102
 Db 4027 ATTGGAACCTTTCTTCTCTCGAGCAATCCGAGCTTGTTCAGCAATGACGACACGCCGTT 4086
 OY 103 CTTATACACCAACTTATACCTAGAAAAGCGCGCTCGATGAT 151
 Db 4087 CGTCGCCAACAACCTTAATCTTAATATACGACGATCGCTCGATTAAT 4135
 RESULT 7
 AAX13487/C
 ID AAX13487 standard; DNA; 9797 BP.
 AC AAX13487;
 XX
 DT 19-MAR-1999 (first entry)
 XX
 DE *Enterococcus faecalis* genome contig SEQ ID NO:550.
 XX
 KW *Enterococcus faecalis*; contig; detection; *Enterococcal* infection;
 KW vaccine; attenuation; computer readable medium; ds.
 XX
 OS *Enterococcus faecalis*.
 XX
 FN M09850555-A2.
 XX
 PD 12-NOV-1998.
 XX
 PE 04-MAY-1998; 98WO-US08985.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Dillon PJ, Kunsch CA;
 DR WPI; 1999-045171/04.
 XX
 PT New isolated *Enterococcus faecalis* polynucleotides and polypeptides
 PT - used to develop products for the detection of *Enterococcus* and for
 PT use in vaccines for prevention or attenuation of *Enterococcal*
 PT infection.
 XX
 PS Claim 1; Page 1805-1810; 2084pp; English.
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the *Enterococcus faecalis* genome.
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the *Enterococcus faecalis* genome with
 CC commercial importance. The products can be used to detect the presence
 CC of *Enterococcus faecalis* in samples. They can also be used for
 CC diagnosing *Enterococcal* infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of *Enterococcus faecalis*, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the *Enterococcus faecalis* nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an *Enterococcal*
 CC infection.

CC specific promoter, preferably senescence associated gene (SAG)-12
 CC promoter connected to a protein-coding DNA sequence not
 CC natively connected to the promoter. The senescence associated
 CC promoters are useful for the creation of transgenic plants with
 CC altered senescence characteristics. Genetic constructs can be
 CC inserted into plants which become effective only upon plant cells
 CC entering senescence. For example, a gene encoding a biosynthetic
 CC enzyme under the control of a senescence-specific promoter can be
 CC inserted into a plant, without having the tissues of the plant
 CC exposed to the excess of cytokinin during pre-senescence growth.
 CC Then at the onset of senescence, the senescence-specific promoter
 CC activates cytokinin production to alter the progression of senescence
 CC in the plant. The present sequence is an alternative version of
 CC a chimeric construct comprising Arabidopsis thaliana SAG12 promoter,
 CC isopenenyl transferase (IPT) gene and NOS-ter sequence.
 CC Note: This sequence is stated to be same as that shown as SEQ ID NO:1
 CC (AA033601) in Fig 3 of the specification. However the sequences differ
 CC at several positions.
 CC XX

Sequence 3182 BP; 1042 A; 526 C; 560 G; 1054 T; 0 other.

Query Match 9.9%; Score 32; DB 24; Length 3182;

Best Local Similarity 50.7%; Pred. No. 3.2; Mismatches 77; Conservative 0; Indels 0; Gaps 0;

DB 120 TCACCTAGAAAAAGCAGCCGCTCGATCGATTCGAAATCTACTCCACCAACCACT 179

DB 424 TCATATATATATACATTCATTCATATATATATATATATATATATATATATAT 365

DB 180 TTGATCTGCTTATGTATATACCAAGTGTCTGATATGATGTGATATATGCTCTT 239

DB 364 ATTCACAGCTTTTCATTTAAAAAACAGTTTCGTATATAAAACTGGAAATATGTTTT 305

DB 240 TCTTCTAGAAATGCTTCCTGCCGATGCTTAT 271

DB 304 AATTATTTGAACGTACATTTGTAAGGTTCT 273

RESULT 12
 AAT42919/C

ID AAT42919 standard; DNA; 3183 BP.

XX AAT42919;

DT 18-JAN-1997 (first entry)

DE SAG12-1 promoter, isopenenyltransferase gene and NOS terminator.

KW SAG12-1 promoter; senescence-associated gene;

KW isopenenyltransferase; NOS terminator; Arabidopsis thaliana;

KW development; gene regulation; 5'-untranslated region;

KW cytokinin biosynthesis; transgenic plant; flowering; seed; fruit;

KW crop improvement; ds.

XX Arabidopsis thaliana.

XX Key

FT promoter

FT promoter

FT promoter

FT promoter

FT terminator

XX WO9629858-A1.

PD 03-OCT-1996.

XX 20-FEB-1996; 96WO-US02313.

XX 29-MAR-1995; 95US-0413135.

PA (MISC) WISCONSIN ALUMNI RES FOUND.

PI Amasino RM, Gan S;

DR WP1; 1996-454877/45.

XX Senescence associated gene promoters, SAG12 and SAG13, - useful for

PT producing genetic constructs for producing transgenic plants having

PT delayed senescence

XX Example: Page 26-27; 38pp; English.

XX The sequence represents a full-length version of the SAG12-1 promoter

CC (uncloned version given in AAT42917) from an Arabidopsis thaliana

CC gene 5'-untranslated region, and has been linked to a cytokinin

CC biosynthesis isopenenyltransferase gene and NOS terminator for

CC senescence-specific gene expression in a transgenic plant. The

CC resulting transgenic plant shows delayed senescence, and shows longer

CC vegetative growth, producing more flowers, seeds or fruit.

XX Sequence 3183 BP; 1042 A; 527 C; 560 G; 1054 T; 0 other;

Query Match 9.9%; Score 32; DB 17; Length 3183;

Best Local Similarity 50.7%; Pred. No. 3.2; Mismatches 77; Conservative 0; Indels 75; Gaps 0;

DB 120 TCACCTAGAAAAAGCAGCCGCTCGATCGATTCGAAATCTACTCCACCAACCACT 179

DB 424 TCATATATATATACATTCATTCATATATATATATATATATATATATATATAT 365

DB 180 TTGATCTGCTTATGTATATACCAAGTGTCTGATATGATGTGATATATGCTCTT 239

DB 364 ATTCACAGCTTTTCATTTAAAAAACAGTTTCGTATATAAAACTGGAAATATGTTTT 305

DB 240 TCTTCTAGAAATGCTTCCTGCCGATGCTTAT 271

DB 304 AATTATTTGAACGTACATTTGTAAGGTTCT 273

RESULT 13
 ABL75912/C

ID ABL75912 standard; cDNA; 285 BP.

XX ABL75912;

DT 14-MAY-2002 (first entry)

DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:5286.

KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPS;

KW inheritance; characteristic; growth; development; disease resistance;

KW environmental adaptability; quality; yield; molecular marker;

KW multigene trait; plant breeding; corn tassel; gene; ss.

XX Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999, 99US-0294093.

XX 21-APR-1998; 98US-082567P.

XX (LALG/) LALGUDI R V.

XX (ITOL/) ITO L Y.


```

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```

```

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159363.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          9 9%; Score 31.8; DB 21; Length 1894;
Best Local Similarity 67.2%; Pred. No. 3;
Matches 45; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      14  GCGGTCACATGCAACGCTAAACAGTCGACGATCTTCTCTCGCAATCCG 73
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       133  GCGGTCACATGCAACGCTAAACAGTCGACGATCTTCTCTCGCTTCT 74
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      74  GCGGTCG 80
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db       73  GCGGCGG 67

RESULT 15
AAK59121/c
ID      AAK59121 standard; cDNA; 1947 BP.
XX      AC      AAK59121.
XX      DT      31-AUG-1999 (first entry)
XX      DE      Maize UDP-glucose dehydrogenase cDNA clone zmdupgh2.
XX      KW      UDP-glucose dehydrogenase; UDPgdH; maize; corn; transgenic plant;
XX      KM      starch; disease resistance; pest resistance, ss.
XX      OS      Zea mays.
XX      FT      Key
XX      FT      CDS      location/Qualifiers
XX      FT      FT      174..1616
XX      FT      FT      /*lag-a
XX      FT      FT      /EC_number="1.1.1.22"
XX      PN      MO9929875-A2.
XX      PD      17-JUN-1999.
XX      PF      09-DEC-1998; 98MO-US26190.
XX      PR      10-DEC-1997; 97US-0987367.

```

XX (PION-) PIONEER HI-BRED INT INC.

PA Bruce WR, Dhuga KS, Lu G, Nichols SE, Saunders CA;
PI Sims LE, Singletary GW, Witcher D, Zhong G;

XX WPI: 1999-385615/32.
DR P-PSDE; AAY06307.

XX New maize UDP-glucose dehydrogenase useful for increasing the
PT biosynthesis of gums in seeds of transgenic plants

PS Claim 16; Fig 3; 82pp; English.

XX This is the nucleotide sequence of maize UDP-glucose dehydrogenase
CC (UDPgdH) (NC 1.1.1.22) cDNA clone Zmudpgh2, a polymorphic variant
CC of clone Zmudpgh1 (see AAX59120). The 2 clones were isolated by PCR
CC amplification of maize cDNA using M13 primers and identified as
CC UDPgdH sequences using a BLAST program. The invention provides
CC maize UDPgdH, its variants and mutants, especially a mutant in
CC which the Cys-272 residue of UDPgdH is replaced by another amino
CC acid residue. UDPgdH (see AAY06306) catalyses a 2-step reaction in
CC which UDP-D-glucose is converted to UDP-6-aldehydo-D-glucose, which
CC is in turn converted to UDP-D-glucuronic acid. With the mutant
CC enzymes, UDP-6-dehydro-D-glucose is the end product. UDPgdH
CC nucleic acids can be used to produce transgenic plants having
CC altered quality or quantity of starch. Expression of the mutant
CC enzyme provides aldehydic starch. A vector in which a promoter is
CC operatively linked to a UDPgdH nucleic acid is used in claimed
CC methods for increasing the stalk or stem length of plants, for
CC increasing insect and pathogen resistance, for increasing root
CC strength, and for increasing resistance to root worm in plants.
CC A vector is which a promoter is linked to a gene that inhibits
CC UDPgdH enzyme activity (e.g. an antisense or ribozyme sequence) is
CC used in methods for increasing the extractability of starch in corn
CC wet milling, and for increasing the nutritional value of plants
CC or plant seeds.

CC
XX
SO Sequence 1947 BP; 420 A; 566 C; 539 G; 422 T; 0 other;

Query Match 9.9%; Score 31.8; DB 20; length 1947;

Best Local Similarity 67.2%; Pred. No. 3;

Matches 45; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 14 GCGGTCGACTGAACCTTAACAGTGAAGATGATCTTCTCTCTGCAATCGT 73
DB 113 GCGGTCGACTGAACCTTAACAGTGAAGATGATCTTCTCTCTGCAATCGT 73
OY 74 GCGGTGG 80
DB 53 GCGGCGG 47

Search completed: August 18, 2003, 13:23:09
Job time : 216 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 12:28:30; Search time 2026 Seconds
(without alignments)
6501.925 Million cell updates/sec

Title: US-10-058-566-5
Perfect score: 322
Sequence: 1 gaattctgcgtcgcgcgtc..... agtgcctagctgagaacatg 322

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	322	100.0	322	6	AX512960 Sequence
2	322	100.0	322	8	AF366294 Zea mays
3	322	100.0	2541	6	AX512958 Sequence
4	187	58.1	187	6	AX512961 Sequence
5	39.4	12.2	145550	2	AC099403 Oryza sat
6	39.4	12.2	173729	2	AC108757 Oryza sat
7	38.8	12.0	202765	2	BX465192 Oryza sat
8	38.8	12.0	204254	2	BX465854 Oryza sat
9	37.8	11.7	179373	10	AF591606 Mouse DNA
10	36.8	11.4	2000	6	AX655393 Sequence
11	36.2	11.2	192623	2	AC131032 Mus muscu
12	36	11.2	112292	9	AC084297 Homo sapi
13	36	11.2	153774	2	AC021009 Homo sapi
14	35.6	11.1	132635	2	AC007577 Drosophila
15	35.6	11.1	156805	2	AC020006 Drosophila
16	35.6	11.1	199016	3	AC008204 Drosophila
17	35.6	11.1	206741	3	AE003747 Drosophila
18	35.6	11.1	250495	10	AL844873 Mouse DNA
19	35.4	11.0	5933	3	AF246991 Drosophila
20	35.4	11.0	6352	3	AY119520 Drosophila
21	35.4	11.0	38193	2	AC017671 Drosophila
22	35.4	11.0	167977	3	AC010010 Drosophila
23	35.4	11.0	168250	3	AC093121 Drosophila
24	35.4	11.0	300610	3	AE003467 Drosophila
25	35.2	10.9	180189	2	AC018864 Homo sapi
26	35.2	10.9	182855	9	AC009292 Homo sapi
27	35.2	10.9	186304	9	AC090136 Homo sapi
28	35.2	10.9	237705	9	AF215845 Homo sapi
29	35	10.9	43886	2	AC021668 Homo sapi
30	34.8	10.8	171708	9	AC067956 Homo sapi
31	34.4	10.7	24725	3	CEK0108 Caenorhabditis
32	34.4	10.7	141084	10	AC127971 Rattus norvegicus
33	34.4	10.7	166167	10	AF611936 Rattus norvegicus
34	34.4	10.7	180291	10	AC129146 Rattus norvegicus
35	34.4	10.7	248625	2	AC106189 Rattus norvegicus
36	34.4	10.7	248675	2	AL844206 Mus musculus
37	34.4	10.7	310448	2	AC131354 Rattus norvegicus
38	34.2	10.6	171930	2	AC118687 Mus musculus
39	34.2	10.6	208765	2	BX004985 Mus musculus
40	34	10.6	59841	9	HS180M12 Human DNA
41	34	10.6	232610	2	AC123157 Rattus norvegicus
42	34	10.6	232631	2	AC110413 Rattus norvegicus
43	34	10.6	247076	2	AC119394 Rattus norvegicus
44	33.8	10.5	113391	2	AC119985 Mus musculus
45	33.8	10.5	161506	9	AL450105 Human DNA

ALIGNMENTS

```

RESULT 1
AX512960
LOCUS AX512960 322 bp DNA linear PAT 03-OCT-2002
DEFINITION Sequence 5 from Patent WO02063021.
ACCESSION AX512960
VERSION AX512960.1 GI:23504095
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1
Albertsen, M., Fox, T., Huffman, G., and Trimnell, M.
Nucleotide sequence mediating male fertility and method of using

```

Pred. No. is the number of results predicted by chance to have a

same
JOURNAL Patent: W0 02063021-A 5 15-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
SOURCE
1. .322
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 79 a 83 c 69 g 91 t
ORIGIN

Query Match 100.0%; Score 322; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.1e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
QY 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180
DB 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180
QY 131 TGTATCTGCTTACTGATGATCAACCAAGTTGCTGATGATGATGATGATGAT 240
DB 131 TGTATCTGCTTACTGATGATCAACCAAGTTGCTGATGATGATGATGATGAT 240
QY 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGAGTTGGTACATGATCTG 300
DB 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGAGTTGGTACATGATCTG 300
QY 301 CCAAGTGTACTGAGAACATG 322
DB 301 CCAAGTGTACTGAGAACATG 322

MSJOUR 2
LACUS AF366294
DEFINITION Zea mays dihydro-flavanoid reductase-like protein (ms*-bs7) gene,
promoter sequence.
ACCESSION AF366294
VERSION AF366294.1 GI:14030552
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Fox,T.W., Trimnell,M.P. and Albertsen,M.C.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7100 N.W. 62nd Ave. P.O. Box 1004, Johnston, IA
50131-1004, USA

FEATURES
SOURCE
1. .322
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
1. .>322
/gene="ms*-bs7"
/note="dihydro-flavanoid reductase-like protein; anther
specific"
1. .322
/gene="ms*-bs7"
misc_feature
165..174
/gene="ms*-bs7"
/note="Myb binding domain"

TATA_signal 269..272
/gene="ms*-bs7"
BASE COUNT 79 a 83 c 69 g 91 t
ORIGIN

Query Match 100.0%; Score 322; DB 8; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.1e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
QY 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180
DB 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180
QY 131 TGTATCTGCTTACTGATGATCAACCAAGTTGCTGATGATGATGATGATGAT 240
DB 131 TGTATCTGCTTACTGATGATCAACCAAGTTGCTGATGATGATGATGATGAT 240
QY 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGAGTTGGTACATGATCTG 300
DB 241 CTCTCTAGAAATGTTCTGCGGATGCTTTATAGAGAGAGTTGGTACATGATCTG 300
QY 301 CCAAGTGTACTGAGAACATG 322
DB 301 CCAAGTGTACTGAGAACATG 322

RESULT 3
AX512958
LOCUS AX512958 2541 bp DNA linear PAT 03-OCT-2002
DEFINITION Sequence 3 from Patent W002063021.
ACCESSION AX512958
VERSION AX512958.1 GI:23504094
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Albertsen,M., Fox,T., Huffman,G. and Trimnell,M.
TITLE Nucleotide sequence mediating male fertility and method of using
same
JOURNAL Patent: W0 02063021-A 3 15-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
SOURCE
1. .2541
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 577 a 693 c 631 g 640 t
ORIGIN

Query Match 100.0%; Score 322; DB 6; Length 2541;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
DB 1 GAATTCCTGCTCGCGCGTCACTGAACCGTAACAGTGAAGTGGATCTCTTCT 60
QY 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
DB 61 CTCTGCAATTCGTGCTGTAAGTAATGAGTGGTCTACTTATCAACCACTTAT 120
QY 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180
DB 121 CACCTAGAAAAGCAGCGCTGCTGATGATGTAATCTACTCCACCAACCCAGCTT 180

```

Db      121 CACGTAGAAAGGACCGCTCTGATGATGCAATCTACTCCACCAACCAAGCTT 180
QY      181 TGTATCTGTTACTGTTGATGACCAAAAGTTGCTATAGATGTCGATTTGCTTT 240
Db      181 TGTATCTGTTACTGTTGATGACCAAAAGTTGCTATAGATGTCGATTTGCTTT 240
QY      241 CTCTCTAGATGTTCTGCTGATGCTTTATAGAGAAGTTGTCACATGATCTCTG 300
Db      241 CTCTCTAGATGTTCTGCTGATGCTTTATAGAGAAGTTGTCACATGATCTCTG 300
QY      301 CCAGTCTAGATGTTCTGCTGATGCTTTATAGAGAAGTTGTCACATGATCTCTG 300
Db      301 CCAGTCTAGATGTTCTGCTGATGCTTTATAGAGAAGTTGTCACATGATCTCTG 300

RESULT 4
AX512961 187 bp DNA linear PAT 03-OCT-2002
LOCUS      AX512961
DEFINITION Sequence 6 from Patent WO02063021.
ACCESSION  AX512961
VERSION     AX512961.1 GI:23504096
KEYWORDS
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
AUTHORS     Albertsen, M., Fox, T., Hoffman, G. and Trimmell, M.
TITLE       Nucleotide sequence mediating male fertility and method of using
            same
JOURNAL     PIONEER HI-BRED INTERNATIONAL, INC. (US)
            Patent: WO 02063021-A 6 15-AUG-2002;
            Location/Qualifiers
            source          1..187
                           /organism="Zea mays"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:4577"
BASE COUNT  42 a 46 c 40 g 59 t
ORIGIN
Query Match      58.1%; Score 187; DB 6; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.2e-49;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      136 CGGTCCTGATGATGCAATCTACCTCCAAACCAACCACTTGTATCTGTTACTG 195
Db      1 CGGTCCTGATGATGCAATCTACCTCCAAACCAACCACTTGTATCTGTTACTG 60
QY      196 TGATCACCAGTTGTGTGATAGATGTCGATTTATGCTTCTCTAGAAATTT 255
Db      61 TGATCACCAGTTGTGTGATAGATGTCGATTTATGCTTCTCTAGAAATTT 120
QY      256 CCGCGGATGCTTTATAGAGAAGTTGTCACATGATCTCTGCACTGTCAGCTGA 315
Db      121 CCGCGGATGCTTTATAGAGAAGTTGTCACATGATCTCTGCACTGTCAGCTGA 180
QY      316 GAACATG 322
Db      181 GAACATG 187

```

```

REFERENCE
AUTHORS     Yun, D.-W., Hahn, J.-H., Yoon, U.-H., Lee, J.-S., Lee, M.-C., Eun, M.Y.
            and Kim, H.-I.
TITLE       Oryza sativa BAC OSJNBa0005C18 genomic sequence
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 145550)
AUTHORS     Hahn, J.-H. and Kim, H.-I.
TITLE       Direct Submission
JOURNAL     Submitted (14-NOV-2001) Rice Genome Sequencing Project, National
            Institute of Agricultural Science and Technology (NIAS), RDA, 249
            Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@da.go.kr,
            Tel:82-31-290-0309, Fax:82-31-290-0308)
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            * 1 86000: contig of 86000 bp in length
            * 86001 86100: gap of unknown length
            * 86101 111963: contig of 25863 bp in length
            * 112063 112063: gap of unknown length
            * 112064 112064: contig of 7620 bp in length
            * 112064 112064: gap of unknown length
            * 119783 119783: gap of unknown length
            * 144017 144017: contig of 24234 bp in length
            * 144018 144117: gap of unknown length
            * 144118 145550: contig of 1433 bp in length.
            * 145550: contig of 1433 bp in length.
            Location/Qualifiers
            source          1..145550
                           /organism="Oryza sativa (japonica cultivar-group)"
                           /mol_type="genomic DNA"
                           /variety="Nipponbare"
                           /db_xref="taxon:39947"
                           /chromosome="9"
                           /clone="OSJNBa0005C18"
BASE COUNT  38466 a 33520 c 33408 g 39755 t 401 others
ORIGIN
Query Match      12.2%; Score 39.4; DB 2; Length 145550;
Best Local Similarity 60.2%; Pred. No. 0.27;
Matches 83; Conservative 0; Mismatches 51; Indels 3; Gaps 1;
QY      156 AATCTACCTCAACCAACCAAGTTGTATCTGTTACTGATCACCACCAAGTTGCTG 215
Db      66598 AATCTACCTCAACCAACCAAGTTGTATCTGTTACTGATCACCACCAAGTTGCTG 215
QY      216 ATACGATGCGATTTATGCTGCTTCTCTAGAAAGTTGTCGATGCTTTATAGA 275
Db      66655 TGATGTACAAATTTCTTGTGCTCTCTCTAGAAAGTTGTCGATGCTTTATAGA 66714
QY      276 GAAGTTGTCGATC 232
Db      66715 GAAGTTGATTTGCTC 66731

RESULT 6
AC108757/c 173729 bp DNA linear HTG 31-JAN-2002
LOCUS      AC108757
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
            OSJNBa0005P13, *** SEQUENCING IN PROGRESS ***, 9 ordered pieces.
ACCESSION  AC108757
VERSION     AC108757.1 GI:18449962
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```


Query Match	Best Local Similarity	54.7%;	Pred. No. 0.91;	Length 179373;
Matches	75;	Conservative	0;	Mismatches 62; Indels 0; Gaps 0;
117	TTATACCTGAGAAACGACGCGCTCTCGATGATGATGCAAAATCTACTCTCAACCAACCCA	176		
116241	TTAGSACCAACACGACGACAGGTTGAGCGCCCTGCTGCTCTACTAAACCAATTC	116300		
177	GCCTTTCATCTGCTTCATCTGCATGACCAAAAGTTGCTGCATGACATGCATTCGATTAATTC	236		
116301	GCAACCTAGCTTTTTCATGTGTGCTCCATCTCTGACCAATATATATGACACTATTTGCA	116366		
237	CT 238			
116361	CT 116362			
RESULT 9	AL591606/c	179373 bp	DNA	linear
LOCUS	Mouse DNA sequence from clone RP23-272014 on chromosome 2, complete			
DEFINITION	sequence.			
ACCESSION	AL591606			
VERSION	AL591606.4			
KEYWORDS	HTG.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Almeida, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk			
COMMENT	Cloned from a library of mouse embryonic fibroblasts (MEFs) from a 19.5 dpc embryo. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30), an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emr, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-272014 is from the RP21-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTER: PHAGE3.6			
FEATURES	Location/Qualifiers			
SOURCE	1..179373			
	/organism="Mus musculus"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:10090"			
	/chromosome="2"			
	/clone="RP23-272014"			
	/clone_1lb="RP21-23"			
BASE COUNT	49498 a 37362 c 39607 g 52706 t			
ORIGIN				

Data from AC105342 was used to finish the clone, AC084297.

repeat__region	/map~"2"
repeat__region	/clone~"RP11-44H9"
repeat__region	/clone_lib~"RCCI-11"
repeat__region	1490..1569
repeat__region	/rpl_family~"MIR"
repeat__region	1594..1785
repeat__region	/rpl_family~"MBR1_type"
repeat__region	2224..2403
repeat__region	/rpl_family~"MBR"
repeat__region	2413..3725
repeat__region	/rpl_family~"MBR"
repeat__region	3735..4157
repeat__region	/rpl_family~"MBR"
repeat__region	5426..5571
repeat__region	/rpl_family~"MIR"
repeat__region	5580..5949
repeat__region	/rpl_family~"BRV1"
repeat__region	7322..7788
repeat__region	/rpl_family~"BRV1"
repeat__region	8080..8284
repeat__region	/rpl_family~"MIR"
repeat__region	8330..8419
repeat__region	/rpl_family~"L1"
repeat__region	8520..8585
repeat__region	/rpl_family~"L1"
repeat__region	8593..8637
repeat__region	/rpl_family~"AT_rich"
repeat__region	11650..11700
repeat__region	/rpl_family~"r-rich"
repeat__region	16745..16642
repeat__region	/rpl_family~"MBR"
repeat__region	19122..19483
repeat__region	/rpl_family~"MBR"
repeat__region	19484..20992
repeat__region	/rpl_family~"MBR"
repeat__region	21000..21539
repeat__region	/rpl_family~"BRV1"
repeat__region	21788..22338
repeat__region	/rpl_family~"L1"
repeat__region	22461..23642
repeat__region	/rpl_family~"L1"
repeat__region	23042..23066
repeat__region	/rpl_family~"AT_rich"
repeat__region	23124..23149
repeat__region	/rpl_family~"AT_rich"
repeat__region	23282..23303
repeat__region	/rpl_family~"AT_rich"
repeat__region	23643..23924
repeat__region	/rpl_family~"Alu"
repeat__region	23925..24550
repeat__region	/rpl_family~"L1"
repeat__region	24637..24657
repeat__region	/rpl_family~"AT_rich"
repeat__region	24669..24750
repeat__region	/rpl_family~"TA)n"
repeat__region	24751..27144
repeat__region	/rpl_family~"L1"
repeat__region	25921..25949
repeat__region	/rpl_family~"AT_rich"
repeat__region	26637..26665
repeat__region	/rpl_family~"(A)n"
repeat__region	26925..26988

Query Match	11.28;	Score 36;	DB 9;	Length 112292;
Best Local Similarity	60.08;	Pred. No. 3.4;		
Matches 60;	Conservative 0;	Mismatches 40;	Indels 0;	Gaps 0;

RESULT 13	AC021009/c	LOCUS	DEFINITION
	AC021009	153774 bp	DNA linear HTG 10-Shp-20000
			Homo sapiens clone RP11-279H15, WORKING DRAFT SEQUENCE, 10
			unordered pieces.

VERSION AC021009.4 GI:10047819
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)

TITLE
JOURNAL
REFERENCE
AUTHORS

Moshireli, A.R., Moshireli, M., Nixon, K., Pacleb, J.M., Park, S.,
Pleiffer, B., Poon, L., Sequeira, A., Sethi, H., Sahr, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierman, L.L. and
Rubin, G.M.
Sequencing of *Drosophila melanogaster*
2 (bases 1 to 132635)
Celniker, S.E., Agbayani, A., Arcaina, T., Baxter, E., Blazej, F.G.,
Bult, C.M., Farnham, D.R., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,
Kam, E., Lee, B., Lewis, S., Li, P., Lomoran, M.A., Mazda, P.,
Moshireli, A.R., Moshireli, M., Nixon, K., Pacleb, J.M., Park, S.,
Pleiffer, B., Poon, L., Sequeira, A., Sethi, H., Sahr, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierman, L.L. and
Rubin, G.M.
Submitted (18-MAY-1999) *Drosophila* Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Aug 2, 1999 this sequence version replaced ql:5629990.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following criteria: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 87 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 807: contig of 807 bp in length
* 808 887: gap of unknown length
* 888 1605: contig of 718 bp in length
* 1606 1685: gap of unknown length
* 1686 2325: contig of 640 bp in length
* 2326 2405: gap of unknown length
* 2406 3474: contig of 1069 bp in length
* 3475 3554: gap of unknown length
* 3555 4424: contig of 870 bp in length
* 4425 4504: gap of unknown length
* 4505 5429: contig of 925 bp in length
* 5430 5509: gap of unknown length
* 5510 6622: contig of 1113 bp in length
* 6623 6702: gap of unknown length
* 6703 7819: contig of 1117 bp in length
* 7820 7899: gap of unknown length
* 7900 9031: contig of 1132 bp in length
* 9032 9111: gap of unknown length
* 9112 9994: contig of 883 bp in length
* 9995 10074: gap of unknown length
* 10075 10863: contig of 789 bp in length
* 10864 10943: gap of unknown length
* 10944 11546: contig of 603 bp in length
* 11547 11626: gap of unknown length
* 11627 12586: contig of 960 bp in length
* 12587 12666: gap of unknown length
* 12667 13832: contig of 1166 bp in length
* 13833 13912: gap of unknown length
* 13913 14559: contig of 647 bp in length
* 14560 14639: gap of unknown length
* 14640 15453: contig of 814 bp in length
* 15454 15533: gap of unknown length
* 15534 16709: contig of 1176 bp in length
* 16710 16789: gap of unknown length
* 16790 17413: contig of 624 bp in length
* 17414 17493: gap of unknown length
* 17494 18143: contig of 650 bp in length
* 18144 18223: gap of unknown length
* 18224 19329: contig of 1106 bp in length
* 19330 19409: gap of unknown length

19410 20351: contig of 942 bp in length
* 20352 20431: gap of unknown length
* 20432 21257: contig of 826 bp in length
* 21258 21337: gap of unknown length
* 21338 22272: contig of 935 bp in length
* 22273 22352: gap of unknown length
* 22353 23492: contig of 1140 bp in length
* 23493 23572: gap of unknown length
* 23573 24323: contig of 750 bp in length
* 24324 24402: gap of unknown length
* 24403 25741: contig of 1333 bp in length
* 25742 25821: gap of unknown length
* 25822 27015: contig of 1194 bp in length
* 27016 27095: gap of unknown length
* 27096 27979: contig of 884 bp in length
* 27980 28059: gap of unknown length
* 28060 29090: contig of 1031 bp in length
* 29091 29170: gap of unknown length
* 29171 30332: contig of 1162 bp in length
* 30333 30412: gap of unknown length
* 30413 31156: contig of 744 bp in length
* 31157 31236: gap of unknown length
* 31237 31850: contig of 614 bp in length
* 31851 31930: gap of unknown length
* 31931 33225: contig of 1252 bp in length
* 33226 33303: gap of unknown length
* 33303 34422: contig of 1120 bp in length
* 34423 34502: gap of unknown length
* 34503 36539: contig of 2037 bp in length
* 36540 36619: gap of unknown length
* 36620 37591: gap of unknown length
* 37592 37671: contig of 972 bp in length
* 37672 39179: gap of unknown length
* 39180 39259: contig of 1508 bp in length
* 39260 40565: gap of unknown length
* 40566 40645: contig of 1306 bp in length
* 40646 41732: gap of unknown length
* 41733 41812: contig of 1087 bp in length
* 41813 43094: gap of unknown length
* 43094 43174: contig of 1282 bp in length
* 43175 45242: gap of unknown length
* 45243 45322: contig of 2068 bp in length
* 45323 46265: gap of unknown length
* 46266 46345: contig of 943 bp in length
* 46346 48317: gap of unknown length
* 48318 48397: contig of 1972 bp in length
* 48398 49996: gap of unknown length
* 49997 50076: contig of 1599 bp in length
* 50077 51343: gap of unknown length
* 51344 51423: contig of 1267 bp in length
* 51424 53538: gap of unknown length
* 53539 53618: contig of 2115 bp in length
* 53619 56731: gap of unknown length
* 56732 56811: contig of 3113 bp in length
* 56812 58914: gap of unknown length
* 58915 58994: contig of 2103 bp in length
* 58995 60245: gap of unknown length
* 60246 60325: contig of 1251 bp in length
* 60326 60326: gap of unknown length
* 60326 62640: contig of 2315 bp in length
* 62641 62720: gap of unknown length
* 62721 65653: contig of 2933 bp in length
* 65654 65733: gap of unknown length
* 65734 67353: contig of 1620 bp in length
* 67354 67433: gap of unknown length
* 67434 70394: contig of 2961 bp in length
* 70395 70474: gap of unknown length
* 70475 73761: contig of 3287 bp in length
* 73762 73841: gap of unknown length
* 73842 76601: contig of 2760 bp in length
* 76602 76681: gap of unknown length
* 76682 80374: contig of 3693 bp in length
* 80375 80454: gap of unknown length
* 80455 86559: contig of 6105 bp in length

*	86560	86639:	gap of unknown length
*	86640	90202:	contig of 3563 bp in length
*	90203	90282:	gap of unknown length
*	90283	94550:	contig of 4268 bp in length
*	94551	94630:	gap of unknown length
*	94631	98832:	contig of 4202 bp in length
*	98833	98912:	gap of unknown length
*	98913	102679:	contig of 3767 bp in length
*	102680	102759:	gap of unknown length
*	102760	107186:	contig of 4427 bp in length
*	107187	107266:	gap of unknown length
*	107267	114485:	contig of 7219 bp in length
*	114486	114565:	gap of unknown length
*	114566	115437:	contig of 872 bp in length
*	115438	115517:	gap of unknown length
*	115518	116166:	contig of 649 bp in length
*	116167	116246:	gap of unknown length
*	116247	117050:	contig of 804 bp in length
*	117051	117130:	gap of unknown length
*	117131	117727:	contig of 597 bp in length
*	117728	117807:	gap of unknown length
*	117808	118625:	contig of 818 bp in length
*	118626	118705:	gap of unknown length
*	118706	119280:	contig of 575 bp in length
*	119281	119360:	gap of unknown length
*	119361	120042:	contig of 682 bp in length
*	120043	120122:	gap of unknown length
*	120123	120575:	contig of 453 bp in length
*	120576	120655:	gap of unknown length
*	120656	121366:	contig of 671 bp in length
*	121367	121406:	gap of unknown length
*	121407	122063:	contig of 657 bp in length
*	122064	122143:	gap of unknown length
*	122144	123652:	contig of 509 bp in length
*	123653	123782:	gap of unknown length
*	123783	123942:	contig of 760 bp in length

Query Match	11.18;	Score 35.6;	DB 2;	Length 132635;
Best Local Similarity	51.28;	Pred. No. 4.6;		
Matches 83;	Conservative	0;	Mismatches 79;	Indels 0;
			Gaps	0;

QY	154	CAAAATCTACGTCACCAACCCAGCTTATGCTGCTACTGTGATACCAAAAGTTGGC	213
Db	21246	CATTCTACTTAAAGGACGACAAAGTAATATCGGTCATGAGTCTCATTCGCAAGATGTA	21187
QY	214	TGATACGATGTCGGCATTAATGCTCTCTCTCTCTAGACATGCTCCGCGCATGCTTATAA	273
Db	21186	TCACGAGAAATACCAACAAATGAGATTTCCTCATACACATTTTGGGAGTGGGACGCTAATA	21127
QY	274	GAGAGGTTGGTCAGCATGATCTCTGCCAGTGCTAGCTGA	315
Db	21126	GCGATCGCTTGCAATTCATACCGACACGTCACATGCGCTTTGGACAGA	21085

RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AC020006	156905 bp	DNA	linear	HTG_03-JAN-2000		
	AC020006	Protophila melanogaster, *** SEQUENCING IN PROGRESS ***.					
	AC020006						
	AC020006.1	GI:6664891					
	HTG: HTGS_PHASE2.						
	Protophila melanogaster (fruit fly)						
	Protophila melanogaster						

REFERENCE	
AUTHORS	1 (bases 1 to 156905)
	Adams, M. and Venter, J.C.

TITLE	Direct Submission
JOURNAL	Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT	This sequence was identified as CDM:10211697 by the submitter for more information on this record e-mail to llj@celera.com

```

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  location/Qualifiers
    1..156905
format="Bioscience Resource Project"
FEATURES
  source
    format="Bioscience Resource Project"

```

```

/organism="Drosophila melanogaster
/mol_type="genomic DNA"
/db_xref="taxon:7227"
BASE COUNT 44447 a 34949 c 34561 g 42948 t
ORIGIN

```

Query Match	11.1%;	Score 35.6;	DB 2;	Length 156905;
Best Local Similarity	51.2%;	Pred. No. 4.7;		
Matches 83; Conservative	0;	Mismatches 79;	Indels 0;	Gaps 0;

[illegible][illegible]

Dy 274 GAGAGGTTGGTCAGCATCGATCTCTGCCAGTGTCTAGCTGA 315
| | | | | | | | | | | |
Db 126660 GCCATCGTTGCATTCTATCCGACGCTCACTCGCTTGAGAGA 126701

Search completed: August 18, 2003, 13:57:18
Job time : 2031 secs

•

•